

Randomness in coin tosses and last digits of primes

Byung Mook Weon* and Yeseul Kim
*School of Advanced Materials Science and Engineering,
SKKU Advanced Institute of Nanotechnology (SAINT),
Sungkyunkwan University, Suwon 440-746, South Korea.*
(Dated: September 30, 2014)

Randomness is a central concept to statistics and physics. A new statistical analysis provides evidence that tossing coins and finding last digits of prime numbers are identical problems regarding equally likely outcomes. This analysis explains why randomness of equally likely outcomes is valid at large numbers.

Randomness is essential in statistics and in making a fair decision [1–4]. Coin tossing is a basic example of a random phenomenon [2]: by flipping a coin, one believes to choose one randomly between heads and tails. Coin tossing is a simple and fair way of deciding between two arbitrary options [3]. It is commonly assumed that coin tossing is random. For a fair coin, the probability of heads and tails is equal, i.e., $\text{Prob}(\text{heads}) = \text{Prob}(\text{tails}) = 50\%$ as illustrated in Fig. 1. This situation is valid only under a condition that all possible orientations of the coin are equally likely [4]. In fact, real coins spin in three dimensions and have finite thickness, so that coin tossing is a physical phenomenon governed by Newtonian mechanics [1–4]. Making a choice by flipping a coin is still important in quantum mechanical statistics [5, 6]. The randomness in coin tossing or rolling dice is of great interest in physics and statistics [7–12]: coin or dice tossing is commonly believed to be random but can be chaotic in real world [13].

A similar situation appears in distribution of prime numbers. Prime numbers are positive integers larger than 1: they are dividable only by 1 and themselves. All primes except 2 and 5 should end in a last digit (j) of 1, 3, 7, or 9. In mathematics, the last digits are believed (without a proof) to be randomly or evenly distributed when numbers are large enough [14]. If the last digits of prime numbers come out with the same frequency, then the probability of the four last digits would be equal, i.e., $\text{Prob}(j) = 25\%$ as illustrated in Fig. 2. The study of the distribution of prime numbers has fascinated mathematicians and physicists for many centuries [14–18]. The distribution of prime numbers is essential to mathematics as well as physics and biology. Particularly in many disparate natural datasets and mathematical sequences, the leading digit (d) is not uniformly distributed, but instead has a biased probability as $P(d) = \log_{10}(1 + 1/d)$ with $d = 1, 2, \dots, 9$, known as the Benfords law [15–17]. Therefore, finding the distribution of last digits is still important. All primes except 2 and 5 should end in a last digit of 1, 3, 7, or 9. Particularly four last digits are believed (not proven yet) to be randomly or evenly distributed when numbers are large enough.

Here we introduce a new statistical analysis, as illustrated in Fig. 3, and clearly show that tossing coins and finding last digits of primes is intrinsically identical in statistics: they are the same problems of equally likely outcomes.

How is randomness defined?

There are many examples for equally likely outcomes: representatively, coin tossing is believed to occur with a probability of 50% between heads and tails. For repeated experiments with a sample, if its frequency between expected outcomes is equal, we can say: the expected outcome of the sample is *random*. Here we suggest a simple way to define the randomness concerning equally likely outcomes.

The relative frequency of each outcome, $f_i = n_i/N$ where n_i is the frequency and N is the total number of repetition, can be various or complicated according to experiments and conditions. The range of frequency for each experiment, $R = n_i^{\max} - n_i^{\min}$, indicates the value between the maximum frequency and the minimum frequency. In statistics, the range tends to be larger, the larger the size of the sample [19, 20]. This tendency can be described by a power-law scaling as $R \sim N^\alpha$ where $\alpha > 0$. Such a power-law scaling commonly appears in physics [21, 22]. Finally, the range of the relative frequency between equally likely outcomes, $R/N = f_i^{\max} - f_i^{\min}$, should have a simple relation as $R/N \sim N^\beta$, where $\beta = \alpha - 1$ (here $\beta < 0$ because $\alpha < 1$). The statistical constraint of $R/N \sim N^\beta$ ($\beta < 0$) implies that the frequency of each outcome becomes equal ($R/N \rightarrow 0$) as the total number of repetition increases ($N \rightarrow \infty$). Consequently, the condition of $R/N \rightarrow 0$ at $N \rightarrow \infty$ explains why randomness is valid only at large numbers, which is known as the law of large numbers in probability theory.

How random is a coin toss?

To rule out physical and mechanical aspects of tossed coins, we use an online virtual coin toss simulation application (<http://www.virtualcointoss.com>) with an ideal coin of zero thickness, where there is no bias between heads and tails, ensuring the equal probabilities for heads and tails. Our experiments with perfectly thin coins enable us to consider only the statistical features of the coin-tossing problems. We actually conducted all five experiments as illustrated in Fig. 1(b). For each experiment, we counted the relative frequency of heads, denoted $f_i = n_i/N$ where n_i is the frequency of heads and

*Electronic address: bmweon@skku.edu

N is the total number of tosses. Different experiments are illustrated by different colors. As seen here, the relative frequency of heads gradually approach to the ultimately expected value for heads, i.e., $\text{Prob}(\text{heads}) = 50\%$ [toward the dashed line] as the total number of tosses increase, although all individual curves are different in shape.

Here we analyze the range of frequency of heads and tails for each experiment [Fig. 1(c)], denoted $R = n_i^{\max} - n_i^{\min}$, which is the value between the maximum frequency and the minimum frequency. In statistics, it is known well that the range tends to be larger, the larger the size of the sample. The average value for the range of frequency in tossed coins shows that R increases with N by a power-law scaling as $R \sim N^\alpha$ where $\alpha \approx 0.4$ (the error bars come from the standard errors).

Finally, we obtain the range of the relative frequency between heads and tails for each experiment, denoted $R/N = f_i^{\max} - f_i^{\min}$. From a power-law scaling of $R \sim N^\alpha$, we expect a simple relation as $R/N \sim N^\beta$, where $\beta = \alpha - 1$. In fact, we obtain $\beta = -0.6237 (\pm 0.0272)$ for coin tossing (the error bars come from the standard errors) as illustrated in Fig. 3(a). This analysis clearly shows $\text{Prob}(\text{heads}) = \text{Prob}(\text{tails}) = 50\%$ by $R/N \rightarrow 0$ at $N \rightarrow \infty$, indicating statistical evidence of randomness for coin tossing at large numbers, consistent with a common belief about coin tossing [8].

How random is a last digit of prime numbers?

Prime numbers are positive integers larger than 1: they are dividable only by 1 and themselves. All primes except 2 and 5 should end in a last digit ($j = 1, 3, 7$, or 9). In mathematics, the last digits are believed (without a proof) to be randomly or evenly distributed when numbers are large enough. If the last digits of prime numbers come out with the same frequency, then the probability of the four last digits would be equal, i.e., $\text{Prob}(j) = 25\%$ [Fig. 2(a)]. From all primes in base 10 for integers up to 10^5 , we conducted the statistical analysis that we did for the above coin tossing. As seen in Fig. 2(b), the relative frequency of last digits gradually approach to the ultimately expected value for heads, i.e., $\text{Prob}(j) = 25\%$ [toward the dashed line]. The range of frequency among

1, 3, 7, and 9 [Fig. 2(c)] increases with the total number of primes with a power-law scaling as $R \sim N^{0.4}$, which is consistent with the case of coin tossing [Fig. 1(c)].

The range of the relative frequency among 1, 3, 7, and 9 shows $R/N \sim N^\beta$ with $\beta = 0.5980 (\pm 0.0164)$ for last digits [Fig. 3(b)], which is completely identical to the case of coin tossing [Fig. 3(a)]. This analysis clearly shows $\text{Prob}(j) = 25\%$ by $R/N \rightarrow 0$ at $N \rightarrow \infty$, indicating that the last digit of primes would occur with the same frequency. This result provides the statistical evidence of randomness for last digits of primes at large numbers.

In summary, we introduce a simple expression for randomness at large numbers. From statistical analyses of coin-tosses and last-digits of primes, we show that the range of the relative frequency between equally likely outcomes (R/N) decreases as the total repetition number (N) increases. A power-law scaling for both cases is found as $R/N \sim N^\beta$ ($\beta \approx 0.6$), implying that the frequency of each outcome becomes equal ($R/N \rightarrow 0$) as the total number of repetition increases ($N \rightarrow \infty$). The condition of $R/N \rightarrow 0$ at $N \rightarrow \infty$ explains why randomness is valid only at large numbers. This result consequently supports that finding last digits of primes is intrinsically identical to tossing coins in statistics both cases are the same problems of equally likely outcomes. Finally our finding the power-law relation between the range of the relative frequency among equally likely outcomes and the total number of repetition would be significant to understand why randomness is valid at large numbers (as known as the law of large numbers). Our finding would be important in statistics, physics, and mathematics.

*This work (NRF-2013R1A2A04008115) was supported by Mid-career Researcher Program through NRF grant funded by the MEST (South Korea).

†Author contributions: B.M.W. conceived of the research, designed the project, conducted and analyzed the coin-tossing experiments, and wrote the manuscript. Y.K. compiled and analyzed the prime number data.

‡Competing financial interests: The authors declare no competing financial interests.

-
- [1] Ford, J. How random is a coin toss? *Phys. Today* **36**, 4047 (1983).
 - [2] Diaconis, P., Holmes, S. and Montgomery, R. Dynamical bias in the coin toss. *SIAM Rev.* **49**, 211235 (2007).
 - [3] Strzalko, J. et al. Dynamics of coin tossing is predictable. *Phys. Rep.* **469**, 5992 (2008).
 - [4] Mahadevan, L. and Yong, E. H. Probability, physics, and the coin toss. *Phys. Today* **64**, 6667 (2011).
 - [5] Murphy, T. E. and Roy, R. Chaotic lasers: The worlds fastest dice. *Nature Photonics* **2**, 714715 (2008).
 - [6] Ferrie, C. and Combes, J. How the result if a single coin toss can turn out to be 100 heads. *Phys. Rev. Lett.* **113**, 120404 (2014).
 - [7] OHagan, T. Dicing with the unknown. *Significance* **1**, 132133 (2004).
 - [8] Hayes, B. Quasirandom ramblings. *Am. Scientist* **99**, 282287 (2011).
 - [9] Vulovic, V. Z. and Prange, R. E. Randomness of a true coin toss. *Phys. Rev. A* **33**, 576582 (2008).
 - [10] Nagler, J. and Richter, P. How random is dice tossing? *Phys. Rev. E* **78**, 036207 (2008).
 - [11] Strzalko, J., Grabski, J., Stefanski, A. and Kapitaniak, T. Can the dice be fair by dynamics? *Int. J. Bifurcation Chaos* **20**, 1175 (2010).
 - [12] Kapitaniak, M., Strzalko, J., Grabski, J. and Kapitaniak, T. The three-dimensional dynamics of the die throw. *Chaos* **22**, 047504 (2012).
 - [13] Le Bellac, M. The role of probabilities in physics. *Prog.*

Biophys. Mol. Biol. **110**, 97105 (2012).

- [14] Granville, A. and Martin, G. Prime number races. *Am. Math. Monthly* **113**, 133 (2006).
- [15] Pain, J.-C. Benfords law and complex atomic spectra. *Phys. Rev. E* **77**, 012102 (2008).
- [16] Luque, B. and Lacasa, L. The first-digit frequencies of prime numbers and Riemann zeta zeros. *Proc. R. Soc. A* **465**, 21972216 (2009).
- [17] Shao, L. and Ma, B.-Q. The significant digit law in statistical physics. *Physica A* **389**, 31093116 (2010).
- [18] Tao, T. Structure and Randomness in the Prime Numbers. in *An Invitation to Mathematics*, Schleicher, D. and Lackmann, M. (eds.) (Springer-Verlag: Berlin Heidelberg, 2011).
- [19] Hozo, S. P., Djulbegovic, B. and Hozo, I. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Med. Res. Methodol.* **5**, 13 (2005).
- [20] Wan, X., Wang, W., Liu, J. and Tong, T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. (arXiv:1407.8038).
- [21] Newman, M. E. J. Power laws, Pareto distributions and Zipfs law. *Contemp. Phys.* **46**, 323351 (2005).
- [22] Frank, S. A. The common patterns of nature. *J. Evol. Biol.* **22**, 15631585 (2009).

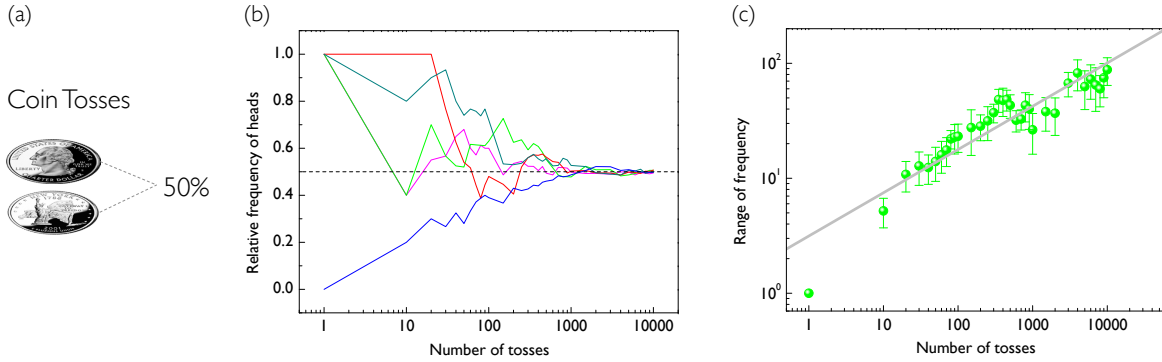


FIG. 1: **Coin tosses.** (a) Schematic illustration of a fair coin with two equally likely outcomes: heads or tails they equally have 50% in probability. (b) The relative frequency of heads taken from five experiments (tossing each coin up to 10^4 repetitions). Different experiments are illustrated by different colors. (c) The range of frequency of heads and tails for each experiment. Here the average value for the range of frequency in tossed coins shows that R increases with N by a power-law scaling as $R \sim N^{0.4}$ (the error bars come from the standard errors).

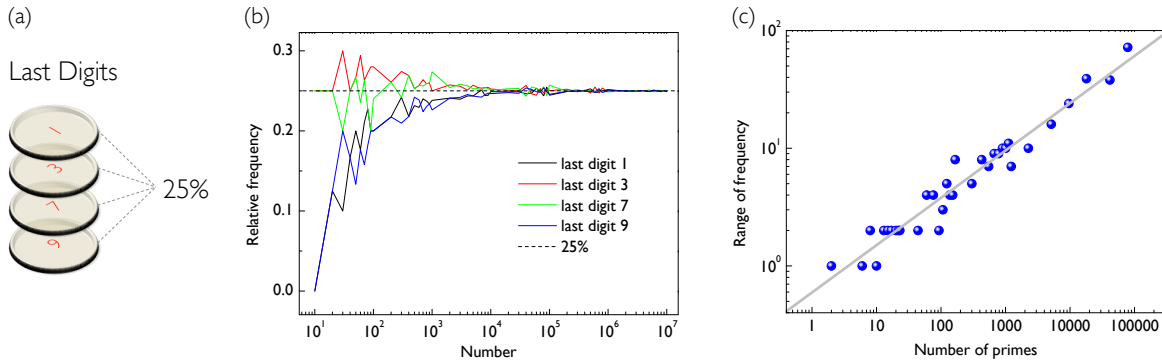


FIG. 2: **Last digits of primes.** (a) Schematic illustration of last digits (1, 3, 7, and 9) of prime numbers up to 10^5 . The probability of each last digit would be equal to be 25%. (b) The relative frequency of last digits gradually approaches to 25% [toward the dashed line]. (c) The range of frequency among last digits increases with the total number of primes with a power-law scaling as $R \sim N^{0.4}$.

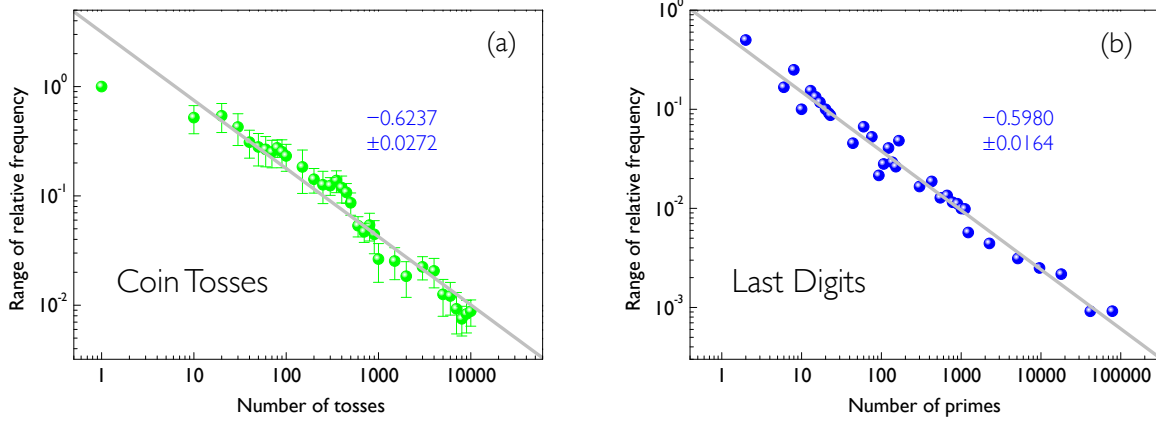


FIG. 3: **Analogy between coin tosses and last digits.** The range of the relative frequency (R/N): (a) for coin tosses between heads and tails and (b) for last digits among four last digits. In both cases, a power-law scaling of $R/N \sim N^\beta$ is found with $\beta = 0.6237 (\pm 0.0272)$ for coin tosses and $\beta = 0.5980 (\pm 0.0164)$ for last digits. This result explains why the randomness for coin tosses or last digits can be valid only at large numbers.